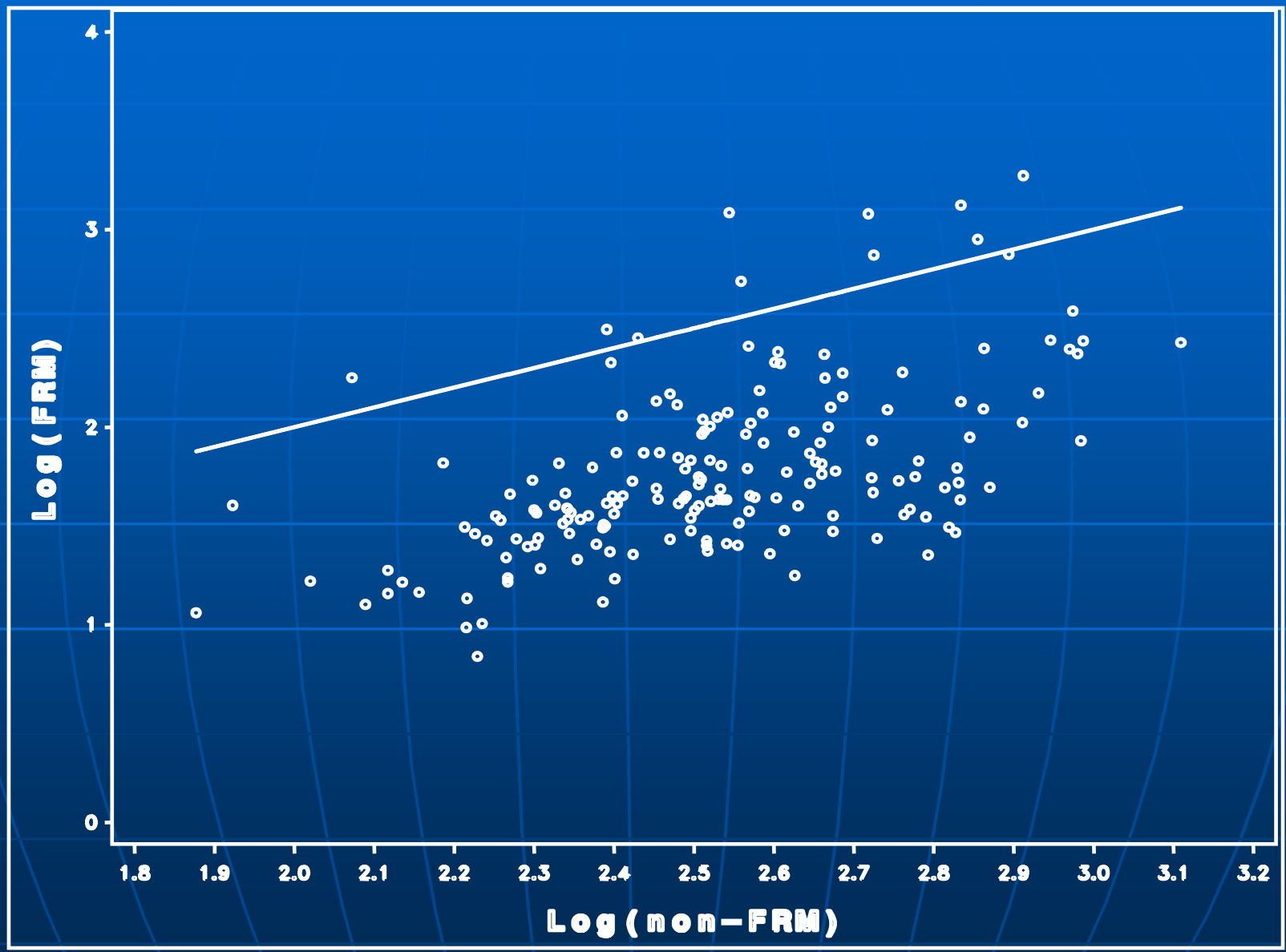
# Describing the Relationship between Continuous TEOM Monitors and Federal Reference Method Monitors When Measuring $PM_{2.5}$

David Wendt, Heather Wiseman, and Steven Bortnick

**Battelle**

# The Issue

- Public need for near real-time PM2.5 AQI reporting and forecasting in MSA's
- Continuous PM2.5 monitoring technologies satisfy this need.
- Continuous monitors are non-FRM/FEM
- FRM/FEM monitors may yield different PM2.5 measurements.

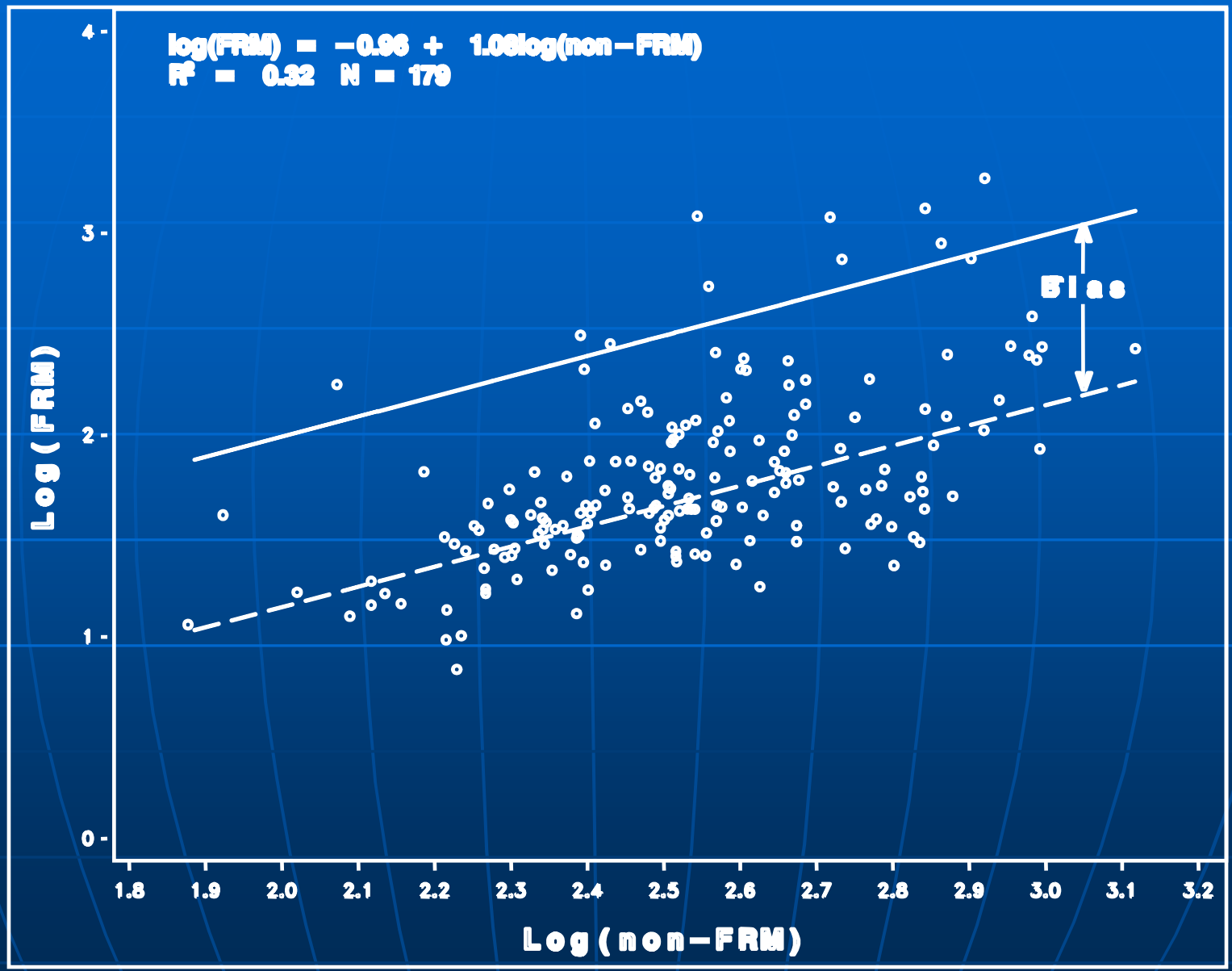

Log (FRM)
4
3
2
1
0
1.8
1.9
2.0
2.1
2.2
2.3
2.4
2.5
2.6
2.7
2.8
2.9
3.0
3.1
3.2
Log (non-FRM)

# The Requirement

- 40 CFR 58 allows for use of non-FRM/FEM technologies in AQI reporting, **BUT** …
- … due to method differences, CFR requires a statistical linear relationship be established between FRM/FEM and non-FRM/FEM measurements.

# The Solution

- Develop linear regression model to relate different PM2.5 measurements, **AND** …
- … use Data Quality Objectives (DQOs) process to ensure acceptable level of regression model performance.


log(FRM) = −0.96 + 1.06log(non−FRM)
R² = 0.32 N = 179
Bias
Log (FRM)
Log (non−FRM)
4
3
2
1
0
1.8
1.9
2.0
2.1
2.2
2.3
2.4
2.5
2.6
2.7
2.8
2.9
3.0
3.1
3.2

# References

- EPA QA/G-4, "Guidance for the Data Quality Objectives Process"
- EPA 454/B-02-002, "Data Quality Objectives (DQOs) for Relating Federal Reference Method (FRM) and Continuous PM2.5 Measurements to Report an Air Quality Index (AQI)"

# Guidance

- In general, EPA 454/B-02-002 suggests a statistical linear regression model is acceptable if:
  - $N \geq 90$, where N is # of measurements used to develop model
  - $R^2 \geq 0.81$, ($R \geq 0.9$) where $R^2$ is measure of model fit
- NOTE: Requirements for N and $R^2$ may vary depending on stringency of DQOs.

# Federal Reference Method (FRM)

- Advantages:
  - "Gold standard" method
  - Homogeneity of monitor type
- Disadvantages:
  - Significant lag time between sample collection and data reporting
  - Costly to operate

# Tapered Element Oscillating Microbalance (TEOM)

- Advantages:
  - Several TEOM monitors in operation
  - Real time concentration reporting
  - Easier to operate
  - Cheaper than FRM
- Disadvantages:
  - Multiple manufacturers
  - Systematically measures lower concentration compared to FRM

# A Caveat

- The composition of $PM_{2.5}$ varies in different geographic areas, which is unobservable
- As a surrogate to unobservable $PM_{2.5}$ composition, meteorological data will be used to better express the relationship between TEOM and FRM monitors.

# New York, New York

# Sites of Interest

- **Bronx County**
  - **I.S. 52**
- Erie County
  - Buffalo
- Essex County
  - Whiteface
- Monroe County
  - Rochester
- Niagara County
  - Niagara Falls
- Queens County
  - Queens College
- Steuben County
  - Pinnacle*

* Continuous TEOM data not available for this site. Thus, Pinnacle is omitted from model building procedures.

# Focus of Analysis:

## I.S. 52

AIRS identification number:
36-005-0110

# Data Sources

- FRM PM2.5 concentrations –

  downloaded from AQS AIRS database by Battelle

- Temperature measurements –

  provided to Battelle by State of New York

- TEOM PM2.5 concentrations –

  downloaded from AQS AIRS database by Battelle

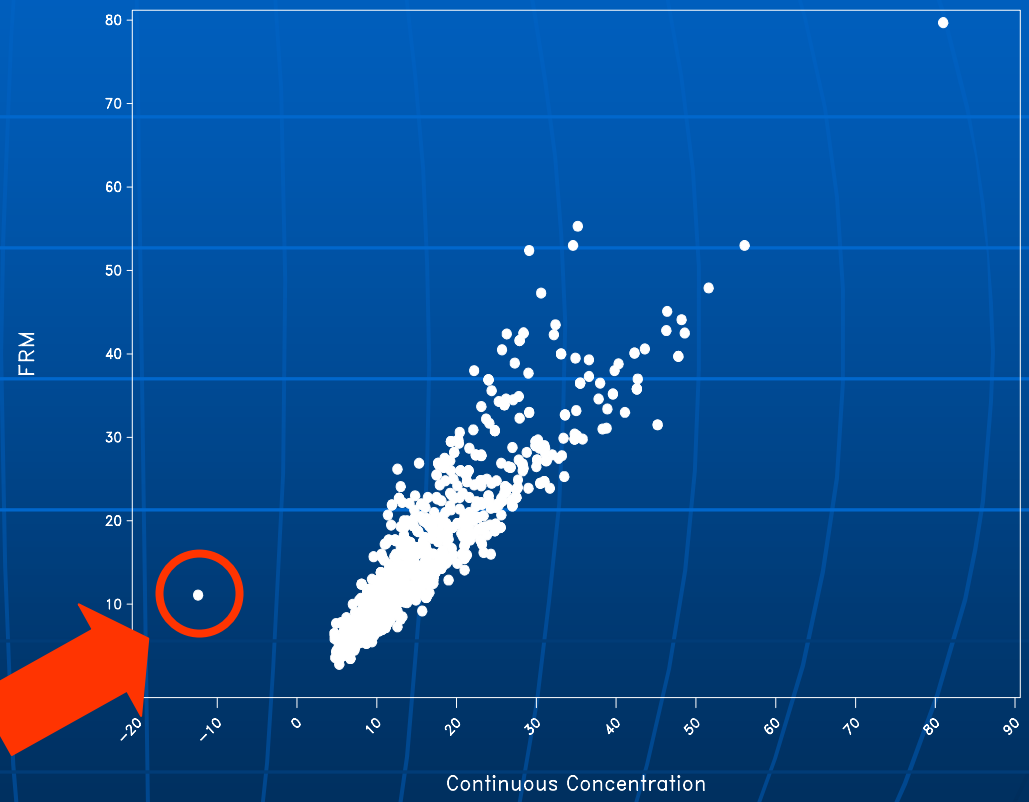
# Basic Model

FRM = Intercept + TEOM +
Temperature +
"Season" + Random Error

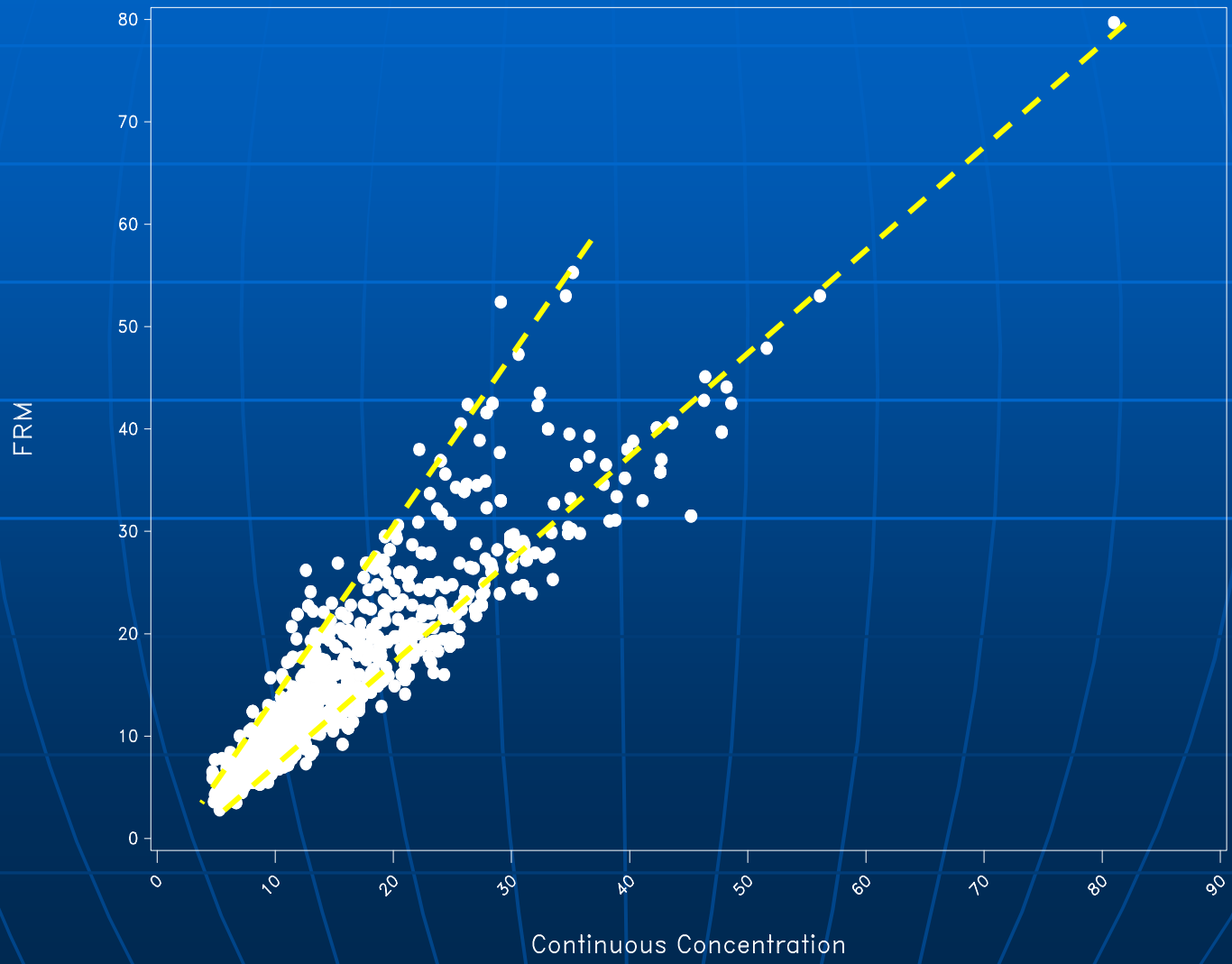
* Goal of $R^2 \geq 0.81$

# One Caveat -

- Exploratory analysis revealed the existence of an outlier for June 23, 2001. Presented is a scatter plot of FRM verses TEOM concentration, including the outlier (lower left).
- This observation was deleted from the data for all analyses.

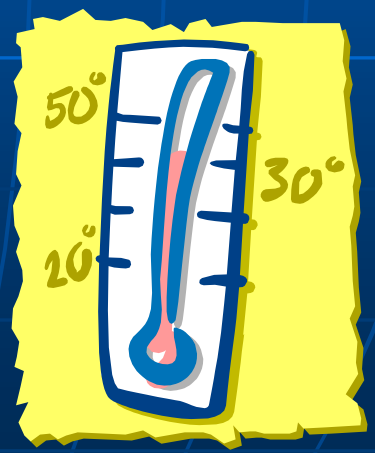


# Plot of FRM verses TEOM Concentrations without Outlier



# Model Building Process

- Separate models were built for each site of interest. Variables were included based on their significance in the models for the individuals site.
- Recall, temperature was the only meteorological variable available for use in analysis.



# Model Building Process (cont.)

- This model contained the following parameters:
  - TEOM concentration
  - TEOM concentration * Temperature interaction
  - Temperature
  - Temperature$^2$
- $R^2$= 0.9143

# Model Building Process (cont.)

- Finally, seasonality was considered as a possible predictor of FRM concentration.
- Two methods for representing time were attempted:
  - Quarterly
    - Winter (December, January, February)
    - Spring (March, April, May)
    - Summer (June July, August)
    - Fall (September, October, November)
  - Monthly
    - 12 calendar months
- Models were constructed using both intervals of time.

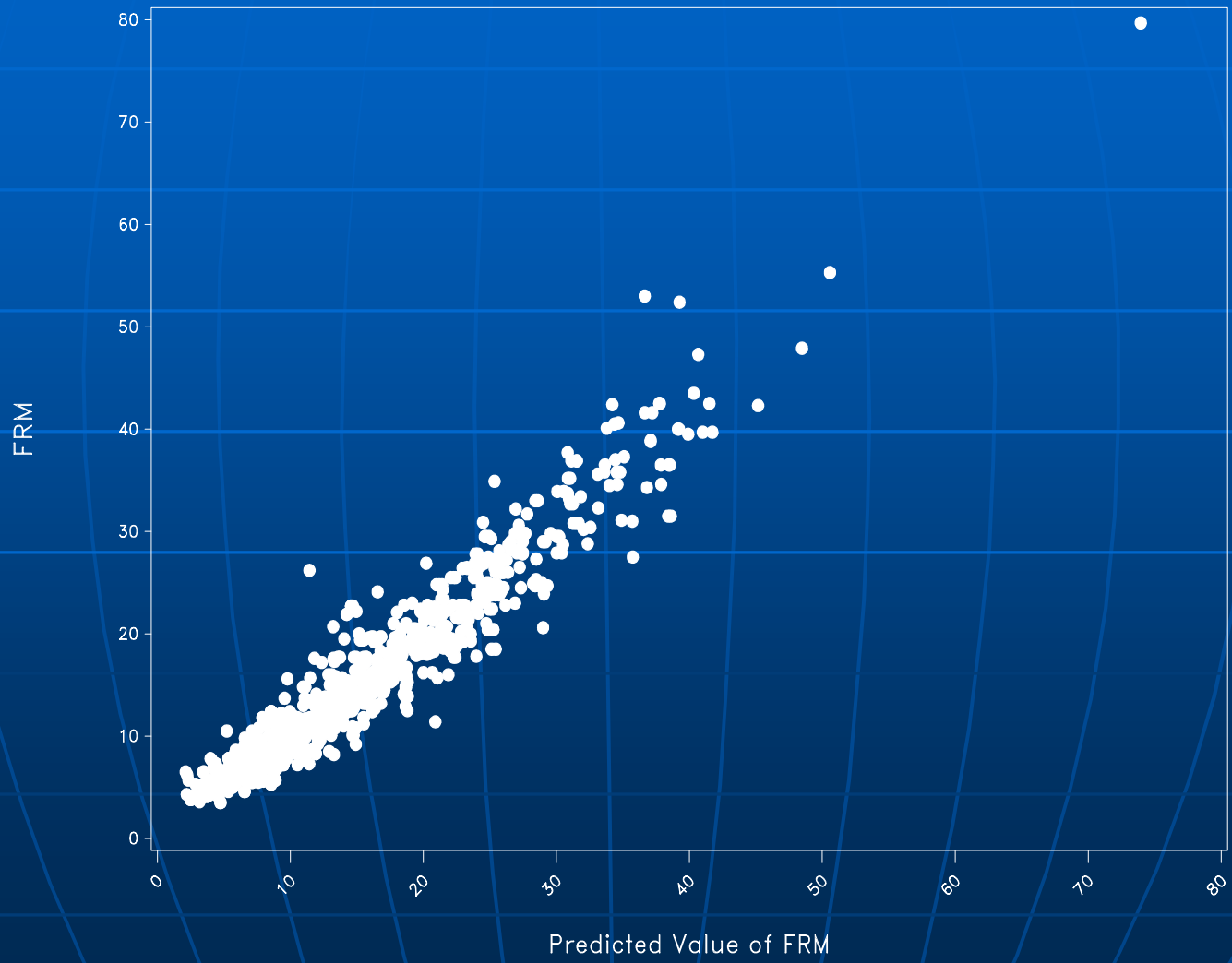
# Model Building Process (cont.)

- Ultimately, time of year expressed as calendar months explained more of the variability in FRM concentration.
- This model contained the following parameters:
  - Eleven calendar month variables

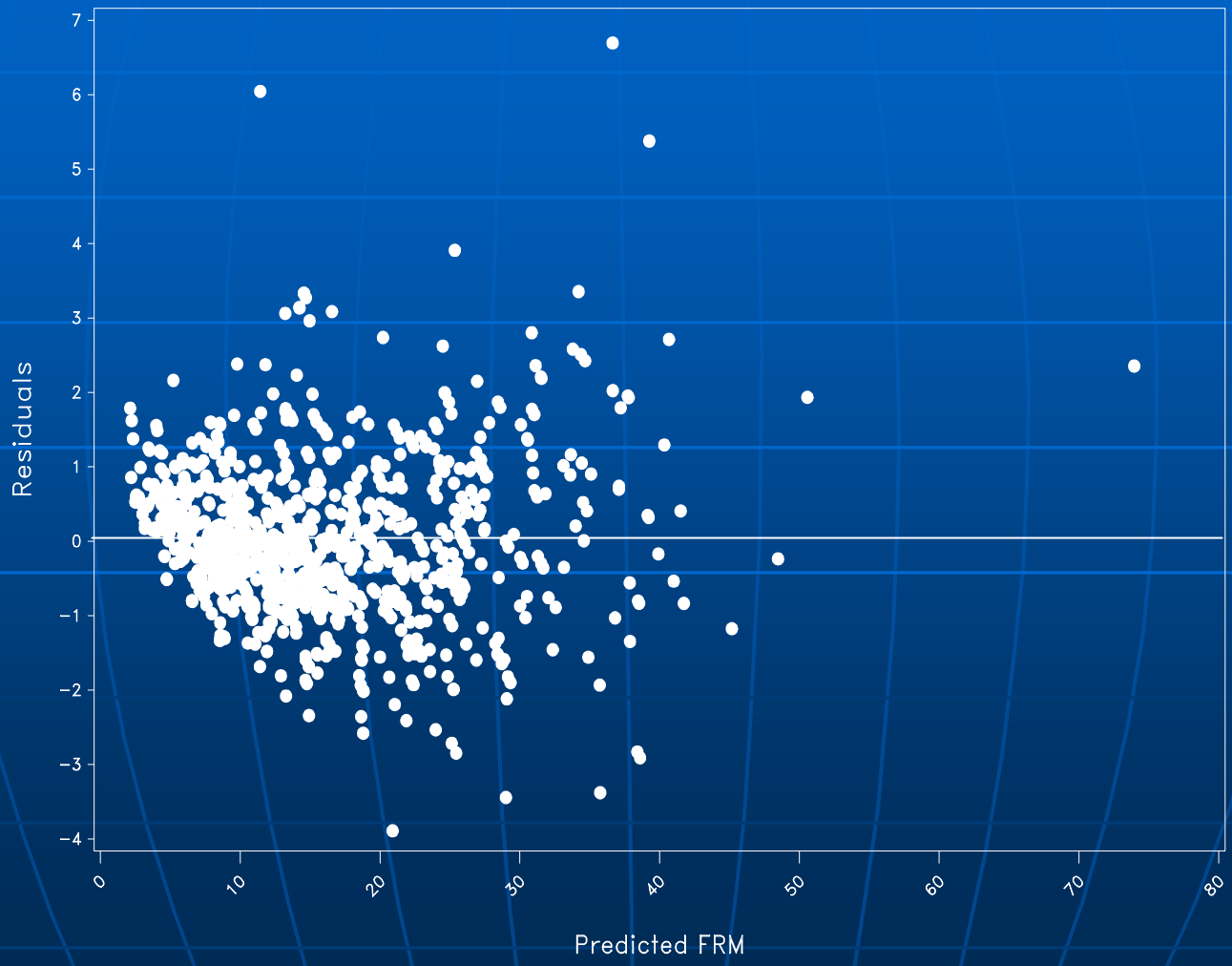
    * January was used as the base line reference month.
  - TEOM concentration
  - Temperature
  - TEOM concentration * Temperature interaction
  - Temperature$^2$
- $R^2$= 0.9211

Small increase from $R^2$ of 0.9143, be careful of over fitting….

# Plot of Observed FRM Values Verses Predicted FRM Values



# Plot of Residuals Verses Fitted Values



# Additional Explorations

- Natural Logarithmic transformations of FRM and/or TEOM concentrations were considered.
- Improvements were minimal, thus the final model does not include transformations.

# Salt Lake City, Utah

# Sites of Interest

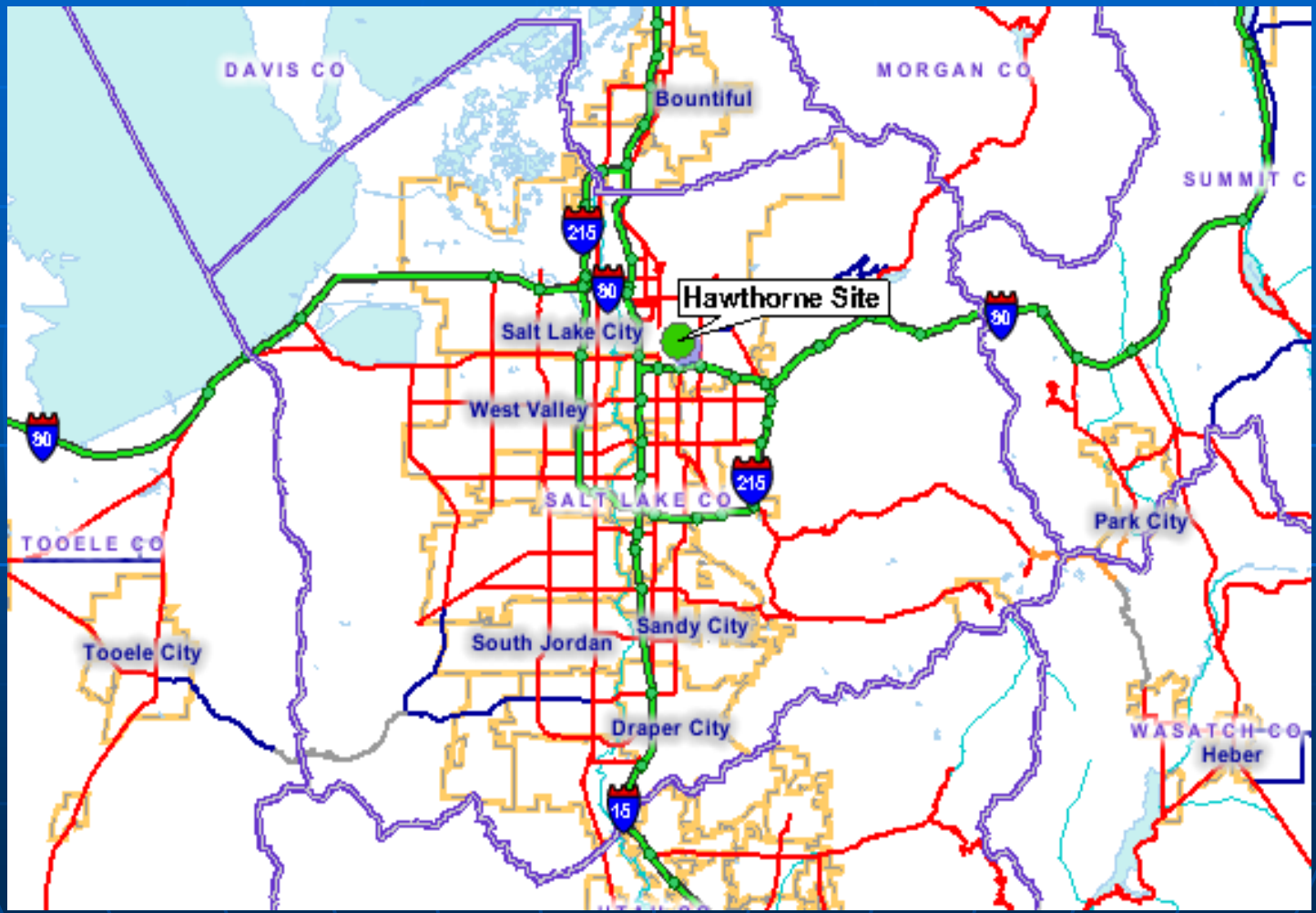
- Cache County
  - Logan
- **Salt Lake County**
  - **Hawthorne**
- Utah County
  - Lindon
- Weber County
  - Harrisville
  - Ogden #2
  - Old Ogden Station

# Focus of Analysis:

## Hawthorne

AIRS identification number:
49-035-3006

# Location of Hawthorne Site




Map from Utah Division of Air Quality, Air Monitoring Center

# Data Sources

- FRM PM2.5 concentrations –

  downloaded from AQS AIRS database by Battelle

- Meteorological measurements (relative humidity, wind direction, wind speed, and temperature) –

  provided to Battelle by Utah Air Monitoring Center

- TEOM PM2.5 concentrations –

  provided to Battelle by Utah Air Monitoring Center

# Data Sources (cont.)

- Data span January 1, 2000 through June 30, 2002
- 819 observation days in which both FRM and TEOM monitors recorded samples.
- One caveat –
  - Three pairs of sample data were removed from analysis due to exceptionally large TEOM samples which appear to be invalid readings.
  - January 22, 2000 (1669.0 $\mu g/m^3$)
    January 23, 2000 (4994.7 $\mu g/m^3$)
    January 24, 2000 (4587.1 $\mu g/m^3$)
- SAS® statistical software was used for all analyses performed

# Basic Model

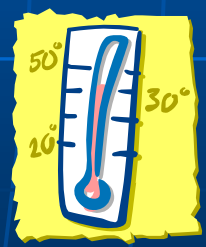
FRM = Intercept + TEOM +
Meteorological Variables +
"Season" + Random Error

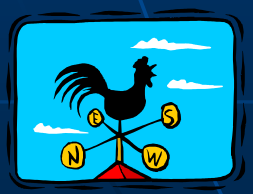
* Goal of $R^2 \geq 0.81$

# Model Building Process

- The first step in the analysis was determining which meteorological variables, in addition to TEOM concentration, were significant main effects. Quadratic forms of the meteorological variables and interactions between TEOM and meteorological variables were also considered.

- Recall, meteorological variables considered in analysis were:
  - relative humidity
  - temperature
  - wind direction
  - wind speed



# Model Building Process (cont.)

- This model contained the following parameters:
  - Wind speed
  - Relative humidity
  - Temperature
  - TEOM concentration
  - TEOM concentration * Relative humidity interaction
  - Relative humidity$^2$
- $R^2$= 0.8063

# Model Building Process (cont.)

- Next, interactions between the meteorological variables were considered for the model.
- This model contained the following parameters:
  - Wind speed
  - Temperature
  - Relative humidity
  - TEOM concentration
  - Relative humidity * Wind speed interaction
  - Relative humidity * Temperature interaction
  - TEOM concentration * Relative humidity interaction
  - Relative humidity$^2$
- $R^2$= 0.8256

# Model Building Process (cont.)

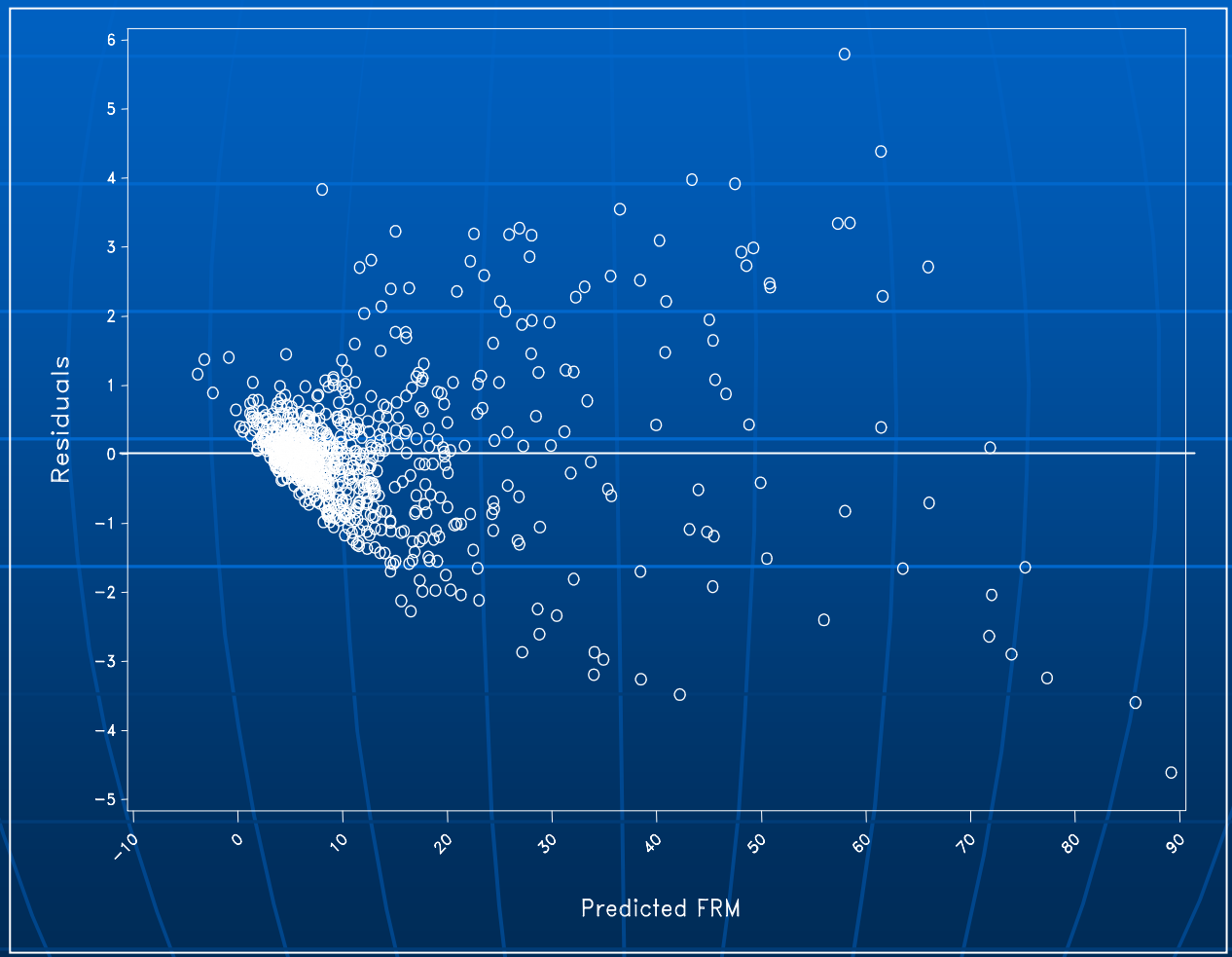
- Finally, seasonality was considered as a possible predictor of FRM concentration.
- Two methods for representing time were attempted:
  - Quarterly
    - Winter (December, January, February)
    - Spring (March, April, May)
    - Summer (June July, August)
    - Fall (September, October, November)
  - Monthly
    - 12 calendar months
- Models were constructed using both intervals of time.

# Model Building Process (cont.)

- Ultimately, time of year expressed as calendar months explained more of the variability in FRM concentration.
- This model contained the following parameters:
  - Wind speed
  - Temperature
  - Relative humidity
  - TEOM concentration
  - Relative humidity * Wind speed interaction
  - Relative humidity * Temperature interaction
  - TEOM concentration * Relative humidity interaction
  - Relative humidity$^2$
  - Eleven calendar month variables

    * January was used as the base line reference month.
- $R^2 = 0.8418$

# Plot of Residuals Verses Fitted Values



# Additional Explorations

- Natural Logarithmic transformations of FRM and/or TEOM concentrations were considered, however, improvements made to the residual plot were minimal, and $R^2$ decreased.
- Log transformations didn't yield additive models which were desirable by Utah. (Average of Hourly Estimates == Daily Average Estimate)

# Notable Points

- It is desirable to create one model for each of the six sites of interest in Utah.
- Ideally, each model would contain the same parameters, differing only in the appropriate coefficients.
- However, all data are not available for all sites.
- Thus, models may differ in the parameters used to describe the relationship between FRM and TEOM monitors.

# Solutions and Future Work

- To deal with missing data-
  - Cluster sites geographically,
  - Build models for clusters of sites instead of individual sites
- Future Work -
  - Explore using breakpoints in relative humidity or other meteorological variables.
  - Explore the use of graphical techniques to identify other transformations.

# Possible Transformation on Wind Speed